# Ways to accept the Properties Existence Constraints in Fuzzy Formal Concept Analysis

Alexander Samoilov
*Samara National Research University*
Samara, Russia
as30606552@gmail.com

Sergey Smirnov
*Institute for the Control of Complex Systems*
*Samara Federal Scientific Center of RAS*
Samara, Russia
smirnov@iccs.ru

***Abstract*—The field of research is the creating of fuzzy concept lattices based on fuzzy data "objects-properties". Our contribution is the account of existential relations on the set of observed and/or measured properties, i.e. "properties existence constraints". Two most well-known approaches to creating of fuzzy concepts lattices are considered: the one-sided threshold and fuzzy closure methods. It is shown that for the more popular one-sided threshold method, potential violations of properties existence constraints in the concept lattice are countered by the rational threshold cut method, previously developed for extracting crisp formal concepts from fuzzy initial data. However, this way is fundamentally unacceptable for the fuzzy closure method. For this case, the idea of special preliminary processing of the initial data is put forward - the "normalization" of the fuzzy set of properties for each object in the training sample. The practical importance of the study is to increase an adequacy of Fuzzy Formal Concept Analysis.**

***Keywords*—*formal context, properties existence constraints, data mining, formal concepts analysis, fuzzy formal concept analysis***

## I. Introduction

One of the most powerful methods of data mining for the last two decades is the Formal Concepts Analysis (FCA) [1-3]. This is the applied branch of the algebraic theory of lattices, which reflects classical representation of the concept [4, 5]. According to this view the concept is the fundamental element of mind that is defined by the extent and intent. The extent is made up by objects, which are applied to the concept. The intent is made up by properties, which are inherent to the concept. These properties are inherent to all objects from the extent of concept.

In FCA intent and extent are associated with the relation $I$ between the set of objects $G$ and set of properties $M$, $I: G \times M \to \{\boldsymbol{True}, \boldsymbol{False}\}$. The tuple $(G, M, I)$ is usually set in the form of the object-properties reflection table and is referred to as a formal context (FC). FC induces Galois operators "↑" and "↓". Formal concept is defined by the bicluster $(X, Y)$, which is formed $X \subseteq G$ (extent) and $Y \subseteq M$ (intent). This bicluster $(X, Y)$ satisfies $X\uparrow = Y$ and $Y\downarrow = X$, where $X\uparrow = \{m \in M | \forall g \in X: I(g, m) = \boldsymbol{True}\}$ and $Y\downarrow = \{g \in G | \forall m \in Y: I(g, m) = \boldsymbol{True}\}$. The set consists of all formal contexts, that is extracted from FC, and is ordered by extent (or intent) inclusion. This set forms the complete lattice and is called the lattice of formal concepts.

On the one hand fuzzy FCA (FFCA) is adaptation to the FCA elasticity of "human concepts" proved by psychologists in the sense that the question of applicability of the concept to the object is the question of degree and not the question of "yes"/"no". People work productively in the conditions of such granulation of opinions [6, 7]. On the other hand, reviewing the realities of accumulating empirical information results in FFCA. In practice, the assessment of the truth of judgments such as "the object $g \in G$ has the property $m \in M$" is often vague. For example, it can be set by an expert or formed by combining information from competing and conflicting sources, and so on. In such cases the truth of judgments also needs to be granulated [8]. Formally both the first and the second are reduced to the fact that $I: G \times M \to L$, where $L$ is a multi-valued truth scale. The actual naming of "fuzzy FCA" is related to the choice of the scale $L = [0,1]$ used in fuzzy logic by L.A. Zadeh.

Another development of FCA is related to the understanding of its hypothetical-deductive nature. In contrast to a posteriori fixed set $G$, the set of measured properties of objects $M$ is formed by the subject of analysis a priori and, as shown in [9], is actually the product of hypotheses about the conceptual description of the research domain of interest. In addition to the composition of the set $M$, existential relations on this set are also defined, known as properties existence constraints (PEC) [10-12]. To derive crisp formal concepts from a fuzzy FC in the presence of a PEC in [12], the method of rational $\alpha$-section of relation $I$ is proposed.

The purpose of this article is to show ways to account for the PEC in the derivation of fuzzy formal concepts.

## II. Properties Existence Constraints

According to [9] subject of Formal Concept Analysis (FCA) a priori forms a system of measured properties, i.e. makes a set $M$ content and two (and only two) existential relations for this set:

- conditionality $\mathrm{C}: M \times M \to \{\boldsymbol{True}, \boldsymbol{False}\}$, when it is established beforehand that having the property $m_j$, every object $g$ necessarily has the property $m_k$ (although the converse may be incorrect), i.e. $\mathrm{C}(m_j, m_k) \leftrightarrow \forall g \in G: m_j \in \{g\}\uparrow \to m_k \in \{g\}\uparrow$. Conditionality reflexive, not symmetric and transitive;
- incompatibility $E: M \times M \to \{\boldsymbol{True}, \boldsymbol{False}\}$, when it is predetermined that having the property $m_j$, every object $g$obviously does not have property $m_k$, and on the contrary, i.e. $E(m_j, m_k) \leftrightarrow \forall g \in G: m_j \in \{g\}\uparrow \to m_k \notin \{g\}\uparrow$. The relation $E$ antireflexive, symmetric and not transitive, but characterized by the so-called "transitivity relative to conditionality", which means $\forall a, b, c \in M: \mathrm{C}(a, b) \wedge E(b, c) \to E(a, c)$.

It should be noted that the relation of conditionality generates a reflexive, symmetrical and transitive relation of inter-conditionality $H$ in a set of properties: $H(m_j, m_k) \leftrightarrow \mathrm{C}(m_j, m_k) \wedge \mathrm{C}(m_k, m_j)$.

Conditionality and incompatibility relations limit set or, in other words, co-existence of properties for training selection objects. According to the subject's a priori hypotheses any object $g \in G$ can only have a "normal" subset of the set of measurable properties $M$ [11]. The subset of measurable properties $Y \subseteq M$ is normal if and only if it is closed and compatible:

- $Y$ closed, if it contains all the properties that are conditioned by any element $Y$, i.e. $\forall m_j \in Y: (\exists m_k \in M: \mathrm{C}(m_j, m_k)) \to m_k \in Y$;
- $Y$ compatible, if any two elements $Y$ are not related by the incompatibility relation, i.e. $\forall m_j \in Y: (\exists m_k \in M: \mathrm{E}(m_j, m_k)) \to m_k \notin Y$.

## III. Two Main Methods for Constructing Lattices of Formal Concepts

To date, the FFCA has developed several methods for constructing lattices of fuzzy formal concepts, as well as their various modifications (see, for example, the review [13], articles [14-19]). The most tested methods include the one-way threshold method [20-22], which is briefly outlined below, and the method using the fuzzy closure operator [23].

### A. The one-way threshold method

The "one-sidedness" of the method is that the fuzzy FC is interpreted asymmetrically as an aggregate of fuzzy sets over the universe $M$, each of which describes one of the FC objects. In other words, each $g \in G$ in the FC is represented as a fuzzy set $\{I(g, m_1)/m_1, I(g, m_2)/m_2, \ldots, I(g, m_n)/m_n\}$, where $n = |M|$, $I(g, m_j)$ – the degree of verity of the assertion "the property $m_j \in M$ is inherent in the object $g$". The usual (crisp) formal concept is defined in this case using the threshold value $\alpha \in [0,1]$ as a pair $(X, Y)$, $X \subseteq G$, $Y \subseteq M$, satisfying the conditions $X\uparrow = Y$ and $Y\downarrow = X$, where $X\uparrow = \{m \in M | \forall g \in X: I(g, m) \geq \alpha\}$ and $Y\downarrow = \{g \in G | \forall m \in Y: I(g, m) \geq \alpha\}$.

Structurally, at the first step of the method, an α-approximation of the fuzzy correspondence $I$ is made, and then a crisp FC $(G, M, I^{(\alpha)})$, where

$$I^{(\alpha)}(g, m) = \begin{cases} \boldsymbol{True}, & \text{if } I(g, m) \geq \alpha; \\ \boldsymbol{False}, & \text{in other cases.} \end{cases}$$

At the second step, the methodological complex of the classical FCA is used to derive crisp concepts from $(G, M, I^{(\alpha)})$ and construct their lattice.

The content of the third step is the transformation of each received crisp concept $(X, Y)$ into a fuzzy one $(X_f, Y)$. The intent $X_f$ is a fuzzy set over the universe $G$. The estimate of the truth $X_f(g)$ of an object $g \in G$ belong to $X_f$ is determined by the degree to which it has all the properties in the content $Y$, or rather, by the assessment of the joint truth (intersection) of fuzzy judgments "to an object $g \in G$ is inherent in the property $m \in Y$" for all properties of $Y$. It is usually proposed to evaluate this degree of membership by a min-conjunction:

$$X_f(g) = \begin{cases} \min_{m \in Y} I(g, m), & \text{if } g \in X; \\ 0, & \text{in other cases.} \end{cases}$$

It is easy to verify that between such fuzzy concepts the same partial order will be preserved as between clear concepts obtained at the intermediate step of the one-way threshold method. If we additionally require $\forall g \in G: Y = \emptyset \to X_f(g) = 1$, we see that the fuzzy concepts constructed in the described way — the biclusters $(X_f, Y)$ with a clear content and fuzzy extent — form a complete lattice of fuzzy concepts.

### B. Fuzzy closure method

The method forms fuzzy concepts with fuzzy extents and fuzzy intents. An algebraic structure is introduced into consideration, called a full residual lattice ("division lattice"), $\boldsymbol{L} = \langle L, \wedge, \vee, \otimes, \to, 0, 1\rangle$ such that:

- $\langle L, \wedge, \vee, 0, 1\rangle$ is a complete lattice with the smallest element 0 and the largest element 1;
- $\langle L, \otimes, 1\rangle$ is a commutative monoid;
- fuzzy conjunction $\otimes$ and fuzzy implication $\to$ satisfy the conjugacy condition $x \otimes y \leq z \leftrightarrow x \leq y \to z$.

The authors of [23] use the Lukasevich operators

$$x \otimes y = \max(x + y - 1, 0),$$

$$x \to y = \min(1 - x + y, 1).$$

Evaluating this formalization, following [24], we note that here, as in other approaches based on fuzzy closure (see, for example, [13, 17]), certain problems arise in interpreting lattices of fuzzy concepts that depend on the algebraic structure used because algebraic operations introduced into the analysis are weakly related to the meaning and pragmatics of applications.

Continuing the review of the fuzzy closure method, we denote by $L^Z$ the set of all fuzzy sets over the universe $Z$.

For fuzzy sets $X \in L^G$ и $Y \in L^M$, fuzzy sets $X\uparrow \in L^M$ и $Y\downarrow \in L^G$ are defined as

$$X\uparrow(m) = \wedge_{g \in G}(X(g) \to I(g, m)),$$

$$Y\downarrow(g) = \wedge_{m \in M}(Y(m) \to I(g, m)).$$

$X\uparrow(m)$ indicates the degree of verity that property $m$ characterizes all objects in a fuzzy set $X$. Similarly, $Y\downarrow(g)$ indicates the verity that all properties in a fuzzy set $Y$ are inherent in object $g$.

A pair $(X, Y) \in L^G \times L^M$ is a fuzzy formal concept if $X\uparrow = Y$ and $Y\downarrow = X$. The set of all fuzzy formal concepts extracted from fuzzy FC is partially ordered by the inclusion of fuzzy intents (or, equally, fuzzy extents) and forms a complete lattice of fuzzy concepts.

According to [23], the discovery of all fuzzy concepts is reduced to calculating all the fixed points of a certain fuzzy closure operator. In a fuzzy FC, the composite operator $\uparrow\downarrow: L^G \to L^G$ is the fuzzy closure operator in $G$, and $\downarrow\uparrow: L^M \to L^M$ is the fuzzy closure operator in $M$. The fixed points of the $\uparrow\downarrow$ and $\downarrow\uparrow$ operators determine the extents and intents of fuzzy formal concepts, respectively.

## IV. Accept the Properties Existence Constraints

It is obvious (see section 2) that for crisp formal concepts extracted from this fuzzy FC the "natural" criterion for accounting for PEC is the normality of sets of properties that determine the content of the constructed concepts [10-12]. When deriving fuzzy formal concepts from a fuzzy FC, this approach needs to be expanded. For a fuzzy concept, the content may be a fuzzy set that is directly incompatible with the crisp normal sets defined by the PEC. Therefore, in this case we should rely on a more common condition for accounting for PEC (the "fundamental" criterion): object generated according to any fuzzy concept of a given fuzzy FC must be characterized by a normal set of properties.

### A. Situation when using the one-side threshold method

It is easy to see that to account for PEC in the one-side threshold method, when α-approximation of fuzzy relation $I$ instead of the standard α-sections use the method of rational $\alpha$-section [12, 25]. In this case the contents of the derived fuzzy formal concepts will become crisp normal subsets of the set of measurable properties $M$ and the "natural" criterion for PEC accounting will be satisfied.

### B. Normalization of formal context for use fuzzy closure method

The method in which the construction of fuzzy concepts uses the fuzzy closure operator does not use threshold values and the contents of the output concepts are fuzzy subsets – defined elements of the set $L^M$. So due to the fact that the PEC requirement is formulated in the language of ordinary sets (objects in the domain of interest can only have normal subsets of measurable properties $M$), there is a need to find a connection between the power set elements $2^M$ and set $L^M$.

It is known that in the case $L = [0,1]$ this connection is established by the fuzzy set decomposition theorem. In our designation for each $g \in G$ we have

$$\{g\}\uparrow = \bigcup_{\alpha\in[0,1]} \alpha \cdot (\{g\}\uparrow)_\alpha, \qquad (1)$$

where $\{g\}\uparrow$ is the fuzzy set of properties of object $g$; $\alpha$ is the threshold value, $\alpha \in [0,1]$, $(\{g\}\uparrow)_\alpha$ is the crisp set of object $g$ properties level $\alpha$.

Now the requirement to carry out the "fundamental" criteria for PEC accounting can be applied to the right side (1): PEC will be carried out if all sets $(\{g\}\uparrow)_\alpha$ are normal. Note that in real FC for each $g \in G$ the number of different "summands" in (1) is finite.

Finally, to account to the PEC we can offer this effective method of preprocessing of fuzzy FC $(G, M, I)$ for constructing the lattice of fuzzy formal concepts:

- according to the available PEC all normal subsets of the set of measured properties $M$ are detected;
- relational $I$ is being normalized: for each $g \in G$ the right part (1) is taken as the initial fuzzy set of properties except for "summands" where the set of properties $(\{g\}\uparrow)_\alpha$ is not normal.

When constructing a formal algorithm that solves the problem of PEC accounting for the case of fuzzy closure, we should take into account some features of the crisp sets that are part of the decomposition (1) and features of the structure of the PEC themselves that allow optimizing the filtering process of normal sets, such as:

- the incompatibility relation defines the presence of incompatible property groups in the set $M$– subsets of pair wise incompatible properties. Similarly, the relation of inter-conditionality defines the groups of inter-conditionality properties;
- when approximation of the fuzzy set $\{g\}\uparrow$ the decision to discard some of the terms in the left part (1) can be made based on the analysis of other terms. If $(\{g\}\uparrow)_{\alpha_1}$ was discarded as a set containing two properties from some group of incompatible properties, then $\forall\alpha_2 < \alpha_1$ the set $(\{g\}\uparrow)_{\alpha_2}$ will be discarded too. If $(\{g\}\uparrow)_{\alpha_1}$ was discarded as a set that does not contain at least one property from some group of inter-conditionality properties then $\forall\alpha_2 > \alpha_1$ the set $(\{g\}\uparrow)_{\alpha_2}$ will be discarded too.

## V. Conclusion

Realized research allow to specify ways to combine the deductive achievements of FFCA in the construction of lattices of formal concepts with understanding the role of the second - a priori - aspect of the common hypothetical-deductive nature of FCA. Specifically, we propose ways to account for existential relations on a set of observed and/or measured properties which provide derivation of correct fuzzy concepts.

A promising task is to explore the influence of properties existence constraints on qualitative and quantitative characteristics fuzzy concept lattices depending on the parameters of these constraints and parameters of initial formal concepts.

Practical value of research results is to improve the adequacy of application of FCA. In particular it is used in the construction of fuzzy formal concepts the important role of which is highly appreciated in different applications (see for example [20, 22, 26]). When clustering incomplete data coming from congruent sources (e.g. in short-term forecasting of traffic flows) - fuzzy FCA can successfully compete with traditional methods such as k nearest neighbors method [27].

## References


[1] B. Ganter and R. Wille, "Formal Concept Analysis. Mathematical foundations," Springer-Verlag Berlin-Heidelberg, 1999

[2] D.I. Ignatov, "Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Filds," Information Retrieval. Revised Selected Papers 8th Russian Summer School, Springer International Publishing, pp. 42-141, 2015.

[3] Formal Concept Analysis Homepage [Online]. URL: http:// www.upriss.org.uk/fca/fca.html.

[4] A.D. Getmanova, "Logics. Advanced course," Moscow Publishing House KNORUS, 2016.

[5] Concept [Online]. URL: https://en.wikipedia.org/wiki/Concept.

[6] S. Pollandt, "Fuzzy-Begriffe: Formale Begriffsanalyse unscharfer Daten," Springer-Verlag Berlin Heidelberg, 1996.

[7] R. Belohlavek and V. Vychodil, "What is a Fuzzy Concept Lattice?" CEUR Workshop Proceedings, vol. 162, pp. 34-45, 2005.

[8] S.V. Smirnov, "Biconstituent phenomenon of information and cognitive data analysis," Procedia Engineering, vol. 201, pp. 773-778, 2017.

[9] D.E. Samoylov, V.A. Semenova and S.V. Smirnov, "Multilevel recursive model of properties existence constraints in machine learning," J. of Physics: Conf. Series, vol. 1096, 012096, 2018.

[10] N. Lammari and E. Metais, "Building and maintaining ontologies: a set of algorithms," Data & Knowledge Engineering, vol. 48, no. 2, pp. 155-176, 2004.

[11] V.A. Pronina and L.B. Shipilina, "Using relationships between attributes to build a domain ontology," Control sciences, vol. 1, pp. 27-32, 2009.

[12] V.P. Ofitserov, V.S. Smirnov, S.V. Smirnov, "Method of the alpha-section of non-strict formal contexts in the formal conceptsanalysis," Proc. of the 16th Int. Conf.Complex Systems:Control and Modeling Problems, pp. 228-244, 2014.

[13] R. Belohlavek, "WhatisaFuzzyConceptLattice? "Rough Sets, Fuzzy Sets, Data Mining and Granular Computing," LNAI, vol. 6743, pp. 19-26, 2011.

[14] R. Belohlavek and V. Vychodil , "Factor Analysis of Incidence Data via Novel Decomposition of Matrices," Int.Conf. on Formal Concept Analysis, LNCS, vol. 5548, pp. 83-97, 2009.

[15] V. Cross, M. Kandasamy and W. Yi, "Comparing Two Approaches to Creating Fuzzy Concept Lattices," Proc.of the North American Fuzzy Information Processing Society, pp. 233-238, 2011.

[16] C.V. Glodeanu, "Attribute Exploration in a Fuzzy Setting," CEUR Workshop Proceedings, vol. 876, pp. 114-128, 2012.

[17] V.V. Pankratieva and S.O. Kuznetsov, "Relations between Proto-Fuzzy Concepts, Crisply Generated Fuzzy Concepts and Pattern Structures," Fundamenta Informaticae, vol. 115, no. 4, pp. 265-277, 2012.

[18] S. Boffa, C. De Maio, A. Di Nola, G. Fenza, A.R. Ferraioli and V. Loia, "Unifying fuzzy concept lattice construction methods," Proc. of IEEE Int. Conf. on Fuzzy Systems, pp. 209-216, 2016.

[19] Z. Zhang, "Constructing L-fuzzy concept lattices without fuzzy Galois closure operation," Fuzzy Sets and Systems, vol. 333, pp.71-86, 2018.

[20] Q.T. Tho, S.C. Hui, A.C.M. Fong and T.H. Cao, "Automatic Fuzzy Ontology Generation for the Semantic Web," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 6, pp. 842-856, 2006.

[21] K.M.Y ang, E.H. Kim, S.H. Hwang and S.H. Choi, "Fuzzy Concept Mining based on Formal Concept Analysis," Int. J. of Computers, vol. 2, no. 3, pp. 279-290, 2008.

[22] C. De Maio, L.V. Fenza, S. Senatore, "Towards Automatic Fuzzy Ontology Generation," Proc. of IEEE Int. Conf. on Fuzzy Systems, pp. 1044-1049, 2009.

[23] R. Belohlavek, B. De Baets, J. Outrata and V. Vychodil, "Computing the lattice of all fixpoints of a fuzzy closure operator," IEEE Trans. on Fuzzy systems, vol. 18, no. 3, pp. 546-557, 2010.

[24] K.E. Wolff, "Position Paper: Pragmatics in Fuzzy Theory," Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, LNAI, vol. 6743, pp. 135-138, 2011.

[25] D.E. Samoylov, V.A. Semenova and S.V Smirnov, "Defuzzification of the initial context in Formal Concept Analysis," CEUR Workshop Proceedings, vol. 2416, pp. 1-9, 2019.

[26] N.G. Yarushkina, V.S. Moshkin, G.R. Ishmuratova, I.A. Andreev and I.A. Moshkina, "Application of fuzzy time series and fuzzy ontology integration in diagnostics of technical systems," Ontology of Designing, vol. 8, no. 4, pp. 594-604, 2018.

[27] A.A. Agafonov, A.S. Yumaganov and V.V. Myasnikov, "Big data analysis in a geoinformatic problem of short-term traffic flow forecasting based on a *k* nearest neighbors metod," Computer Optics, vol. 42, no. 6, pp. 1101-1111, 2018. DOI: 10.18287/2412-6179-2018-42-6-1101-1111.